AMENDMENT TO THE CLAIMS

1. (Original) A computer-readable medium including instructions readable by a computer which, when implemented perform steps comprising:

    generating a speech-based phonetic description of a word without reference to the text of the word;

    generating a text-based phonetic description of the word based on the text of the word;

    aligning the speech-based phonetic description and the text-based phonetic description on a phone-by-phone basis to form a single graph; and

    selecting a phonetic description from the single graph.

2. (Original) The computer-readable medium of claim 1, further comprising generating the speech-based phonetic description based on a user's pronunciation of the word.

3. (Original) The computer-readable medium of claim 2, further comprising decoding a speech signal representing the user's pronunciation of the word to generate the speech-based phonetic description of the word.

4. (Currently Amended) The computer-readable medium of claim 2̶3, wherein decoding a speech signal comprises identifying a sequence of syllable-like units from the speech signal.

5. (Original) The computer-readable medium of claim 4, further comprising generating a set of syllable-like units using mutual information before decoding a speech signal to identify a sequence of syllable-like units.

6. (Original) The computer-readable medium of claim 5, wherein

generating a syllable-like unit using mutual information comprises:

        calculating mutual information values for pairs of sub-word units in a training dictionary;

        selecting a pair of sub-word units based on the mutual information values; and

        merging the selected pair of sub-word units into a syllable-like unit.

7. (Original) The computer-readable medium of claim 2, wherein generating the text-based phonetic description comprises using a letter-to-sound rule.

8. (Original) The computer-readable medium of claim 1, wherein selecting a phonetic description from the single graph comprises comparing a speech sample to acoustic models of phonetic units in the single graph.

9. (Original) A computer-readable medium having computer-executable instructions for performing steps comprising:

        receiving text of a word for which a phonetic pronunciation is to be added to a speech recognition lexicon;

        receiving a representation of a speech signal produced by a person pronouncing the word;

        converting the text of the word into at least one text-based phonetic sequence of phonetic units;

        generating a speech-based phonetic sequence of phonetic units from the representation of the speech signal;

        placing the phonetic units of the at least one text-based phonetic sequence and the speech-based phonetic sequence in a search structure that allows for transitions between phonetic units in the text-based

> phonetic sequence and phonetic units in the speech-based phonetic description; and
>
> selecting a phonetic pronunciation from the search structure.

10. (Original)   The computer-readable medium of claim 9, wherein placing the phonetic units in a search structure comprises aligning the speech-based phonetic sequence and the at least one text-based phonetic sequence to identify phonetic units that are alternatives of each other.

11. (Original)   The computer-readable medium of claim 10, wherein aligning the speech-based phonetic sequence and the at least one text-based phonetic sequence comprises calculating a minimum distance between two phonetic sequences.

12. (Original)   The computer-readable medium of claim 10, wherein selecting the phonetic pronunciation is based in part on a comparison between acoustic models of phonetic units and the representation of the speech signal.

13. (Original)   The computer-readable medium of claim 9, wherein generating a speech-based phonetic sequence of phonetic units comprises:

> generating a plurality of possible phonetic sequences of phonetic units;
>
> using at least one model to generate a probability score for each possible phonetic sequence; and
>
> selecting the possible phonetic sequence with the highest score as the speech-based phonetic sequence of phonetic units.

14. (Original)   The computer-readable medium of claim 13, wherein using at least one model comprises using an acoustic model and a language model.

15. (Original)   The computer-readable medium of claim 14, wherein using a language model comprises using a language model that is based on syllable-like units.

16. (Original)   The computer-readable medium of claim 13, wherein selecting a phonetic pronunciation comprises scoring paths through the search structure based on at least one model.

17. (Original)   The computer-readable medium of claim 16, wherein the at least one model comprises an acoustic model.

18. (Original)   The computer-readable medium of claim 10, wherein the search structure contains a single path for a phonetic unit that is found in both the text-based phonetic sequence and the speech-based phonetic sequence.

19. (Original)   A method for adding an acoustic description of a word to a speech recognition lexicon, the method comprising:

    generating a text-based phonetic description based on the text of a word;

    generating a speech-based phonetic description without reference to the text of the word;

    aligning the text-based phonetic description and the speech based phonetic description in a structure, the structure comprising paths representing phonetic units, at least one path for a phonetic unit from the text-based phonetic description being connected to a path for a phonetic unit from the speech-based phonetic description;

selecting a sequence of paths through the structure; and generating the acoustic description of the word based on the selected sequence of paths.

20. (Original)   The method of claim 19, wherein selecting a sequence of paths comprises generating a score for a path in the structure.

21. (Original)   The method of claim 20, wherein generating a score of a path comprises comparing a user's pronunciation of a word to a model for a phonetic unit in the structure.

22. (Original)   The method of claim 20, further comprising generating a plurality of text-based phonetic descriptions based on the text of the word.

23. (Original)   The method of claim 22, wherein generating the speech-based phonetic description comprises decoding a speech signal comprising a user's pronunciation of the word.

24. (Original)   The method of claim 23, wherein decoding a speech signal comprises using a language model of syllable-like-units.

25. (Original)   The method of claim 24, further comprising constructing the language model of syllable-like units through steps of:

 calculating mutual information values for pairs of syllable-like units in a training dictionary;

 selecting a pair of syllable-like units based on the mutual information values; and

 removing the selected pair and substituting a new syllable-like unit in place of the removed selected

pair in the training dictionary.

26.   (Original)   The method of claim 25, further comprising:
        recalculating mutual information values for remaining
            pairs of syllable-like units in the training
            dictionary;
        selecting a new pair of syllable-like units based on the
            recalculated mutual information values; and
        removing the new pair of syllable-like units and
            substituting a second new syllable-like unit in
            place of the new pair of syllable-like units in the
            training dictionary.

27.   (Original)   The method of claim 26, further comprising
    using the training dictionary to generate a language model of
    syllable-like units.